

Supplementary Material

Integrating Multi-Cohort Machine Learning and Clinical Sample Validation to Explore Peripheral Blood mRNA Diagnostic Biomarkers for Prostate Cancer

Xingyu Zhong^{1,*}, Yuxuan Yang^{1,*}, Haodong He¹, Yifan Xiong¹, Mingliang Zhong¹, Shaogang Wang^{1, #}, Qidong Xia^{1, #}

¹ Department and Institute of Urology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, No.1095Jiefang Avenue, 430030, Wuhan, China.

* These authors contributed equally to this work and share first authorship.

*** Correspondence:**

Shaogang Wang

sgwangtjm@163.com

Qidong Xia

qidongxia_md@163.com

Supplementary Tables

Table S1. Primers used for 9-gene amplification.

Gene	Forward Primer (5' to 3')	Reverse Primer (5' to 3')
JPH4	GTGTGCCCTACCATCAGGC	GAACGGCGAAAGAATCCGC
RASL12	CTGCCCTGACCGTGAAGTTT	TCTCCTCGGAGCTGTAGGTG
AOX1	TGTCGATCCTGAAACAATGCTG	GGTGATGGGGTTGTATCGTGA
SLC18A2	CCTGAATGAAAACGTGCAAGTTG	AGTAGTCCTATGAAAGGGTTGGT
PDZRN4	CGAGTTGTGTCTGTGTTAGCAG	ACATAAATGCCGGTGTCTTCTTC
P2RY2	GGTGTCTGGGCGTCTTACG	TGGTGGTGACAAAGTAGAGCA
B3GNT8	CTGCTCAGATACTGATGTCCCC	CTGCTCAGATACTGATGTCCCC
KCNQ5	AAGCCGCTCTCTTACACGAG	AGAAAAACGAAAGCGTGGTAGAT
APOBEC3C	CTTGGTTCTGCGACGACATAC	TCCTGGTAACATGGATACTGGAA
GAPDH	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA

Table S2. Raw qRT-PCR data from cell lines.

Gene	Cell Line	2 ^{-ΔΔCT}		
		1	2	3
JPH4	RWPE1	1.141	0.878	0.998
	LNCAP	0.268	0.244	0.32
	C42	0.377	0.211	0.371
	DU145	1.047	1.014	0.908
	PC3	1.938	1.523	1.251
	RV1	1.131	1.379	1.239
	RV1	1.131	1.379	1.239
RASL12	RWPE1	0.902	1.24	0.894
	LNCAP	0.176	0.233	0.216
	C42	0.348	0.307	0.271
	DU145	1.075	1.101	1.097
	PC3	0.206	0.229	0.208
	RV1	0.403	0.54	0.444
	RV1	0.403	0.54	0.444
AOX1	RWPE1	0.938	1.119	0.953
	LNCAP	0.174	0.121	0.146
	C42	0.171	0.134	0.107
	DU145	0.662	0.649	0.663
	PC3	0.26	0.279	0.197
	RV1	0.047	0.047	0.041
	RV1	0.047	0.047	0.041
SLC18A2	RWPE1	0.841	1.272	0.935
	LNCAP	0.133	0.111	0.136
	C42	0.203	0.279	0.134
	DU145	3.417	2.674	4.295
	PC3	0.072	0.086	0.058
	RV1	0.467	0.524	0.831
	RV1	0.467	0.524	0.831

PDZRN4	RWPE1	1.072	1.199	0.778
	LNCAP	0.125	0.107	0.143
	C42	2.551	2.651	2.361
	DU145	0.22	0.157	0.203
	PC3	0.91	0.827	0.694
	RV1	6.168	8.405	7.797
P2RY2	RWPE1	1	1.064	0.94
	LNCAP	0.092	0.102	0.093
	C42	0.096	0.12	0.12
	DU145	1.908	1.828	1.507
	PC3	5.869	6.222	5.207
	RV1	0.182	0.19	0.232
B3GNT8	RWPE1	1.037	0.941	1.025
	LNCAP	0.058	0.096	0.105
	C42	0.146	0.121	0.187
	DU145	0.22	0.285	0.254
	PC3	0.419	0.343	0.283
	RV1	0.206	0.216	0.245
KCNQ5	RWPE1	0.994	1.031	0.976
	LNCAP	0.141	0.162	0.197
	C42	0.671	0.762	0.665
	DU145	0.773	0.946	0.905
	PC3	0.132	0.238	0.251
	RV1	9.419	10.994	12.001
APOBEC3C	RWPE1	1.254	0.811	0.983
	LNCAP	0.001	0.002	0.002
	C42	0.003	0.004	0.003
	DU145	0.81	0.94	0.933
	PC3	5.625	4.926	5.096
	RV1	0.006	0.008	0.006

Table S3. Raw qRT-PCR data from plasma samples.

BPH	Gene		PCa	Gene	
	AOX1	B3TNG8		AOX1	B3TNG8
1	3.12	0.732	1	0.64	0.531
2	2.75	0.466	2	0.94	2.855
3	2.16	0.51	3	0.61	0.838
4	2.1	0.589	4	0.62	0.438
5	1.39	2.361	5	0.92	0.217
6	1.57	19.741	6	0.65	0.176
7	2.18	8.325	7	1.23	1.778
8	1.4	1.001	8	0.76	0.329
9	3.88	0.659	9	1.12	0.189

10	16.18	0.463	10	3.52	0.257
11	2.02	0.978	11	0.62	2.563
12	18.61	2.969	12	0.67	1.089
13	1.21	7.867	13	0.26	0.408
14	1.31	13.623	14	0.44	0.27
15	0.97	0.643	15	1.48	0.197
16	0.88	2.957	16	0.52	1.576
17	6.99	2.935	17	0.54	0.35
18	6.08	1.382	18	1	0.178
19	1.22	1.095	19	2.85	0.699
20	11.39	0.998	20	1.82	0.523
21	3.027	3.099	21	0.56	0.583
22	1.418	1.672	22	0.57	0.379
23	3.113	3.188	23	0.21	0.269
24	0.943	1.545	24	1.8	0.919
25	2.288	1.329	25	0.77	1.582
26	4.146	4.949	26	0.18	0.682
27	1.328	1.719	27	0.447	0.443
28	3.190	2.928	28	0.688	0.468
29	10.177	14.104	29	0.929	1.273
30	2.215	2.226	30	2.013	2.693
			31	0.500	0.747
			32	0.673	0.869
			33	1.086	0.924
			34	0.807	1.176
			35	1.328	1.298
			36	1.746	1.441
			37	1.277	2.019
			38	0.967	0.774
			39	1.283	1.412
			40	0.693	0.390
			41	0.699	0.681
			42	3.315	1.822

Supplementary Figures

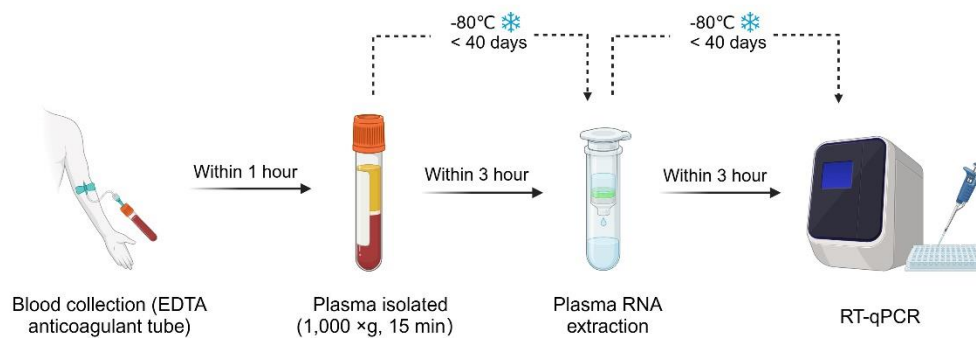


Figure S1. Workflow for the collection and processing of plasma RNA samples.

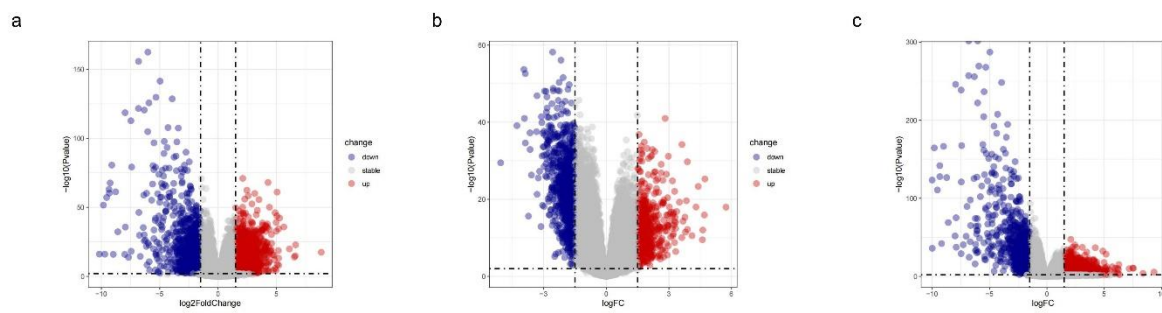


Figure S2. Differentially expressed genes in PCa screened using *DESeq2* (a), *edgeR* (b), and *limma* (c).

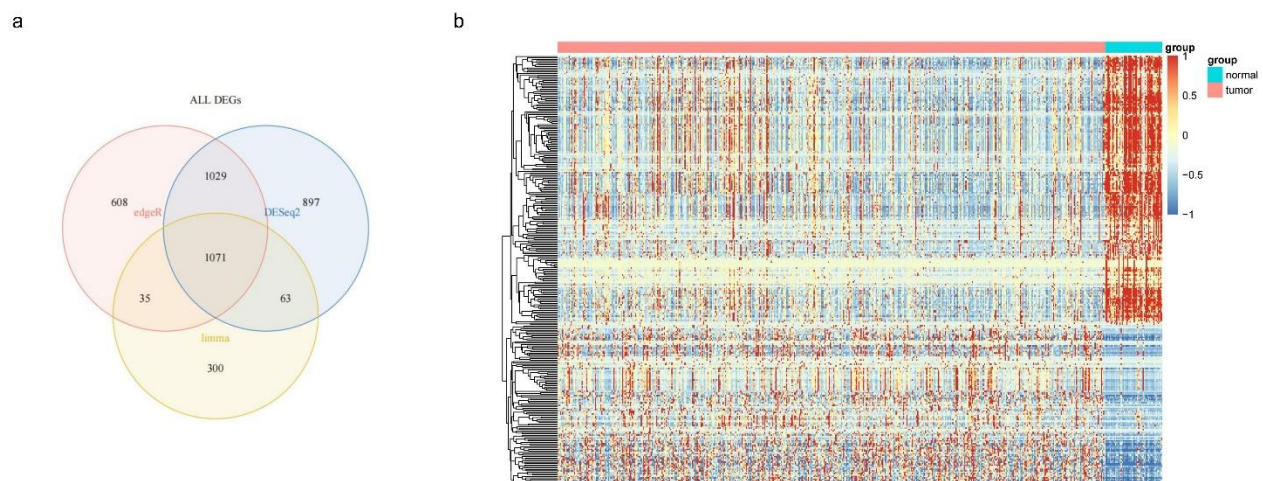


Figure S3. (a) Integration of three machine learning methods to identify 1,071 PCa-associated differentially expressed genes. (b) Heatmap of expression profiles.

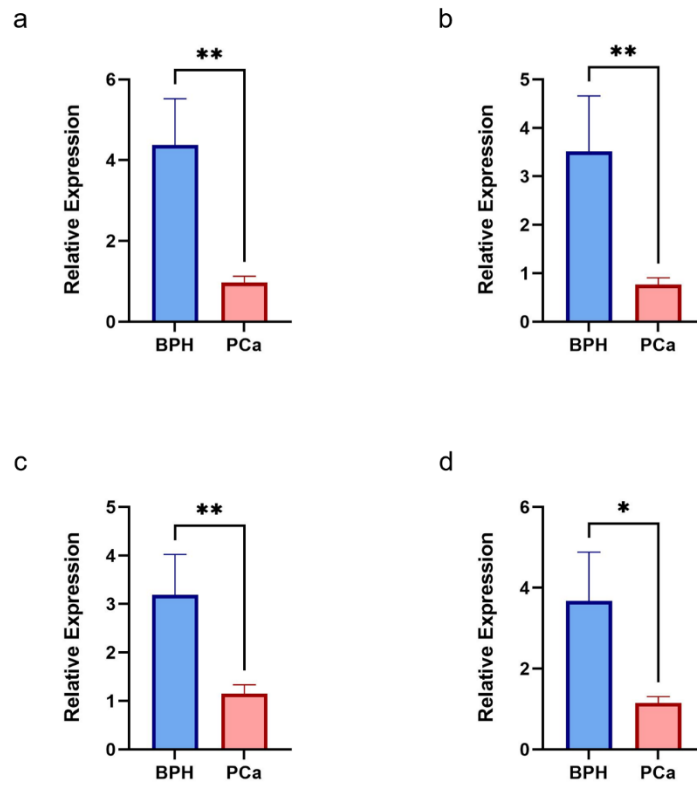


Figure S4. (a-b) Expression profiles of *AOX1* (a) and *B3GNT8* (b) in fresh plasma RNA samples; (c-d) Corresponding expression patterns of *AOX1* (c) and *B3GNT8* (d) in long-term preserved (>30 days) plasma RNA samples. * $p < 0.05$, ** $p < 0.01$.

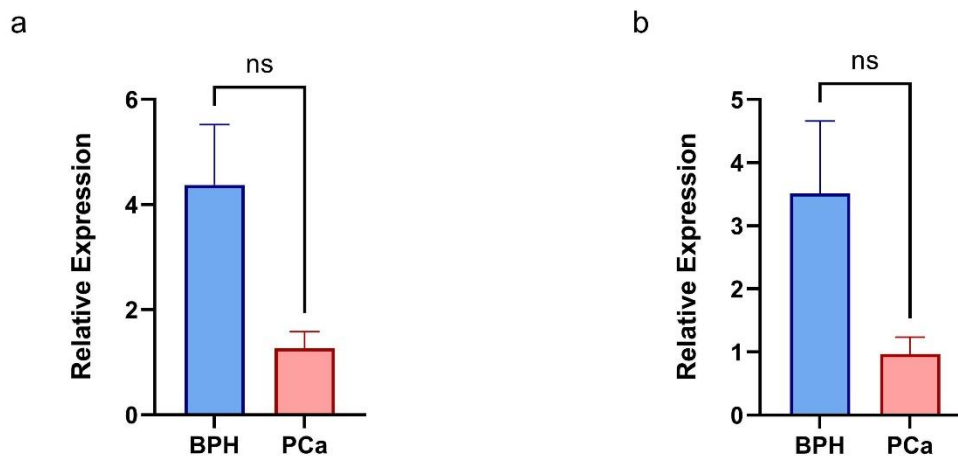


Figure S5. Expression of *AOX1* (a) and *B3GNT8* (b) in plasma from BPH patients and PCa patients with ISUP grade ≤ 2 . ns: no significance.

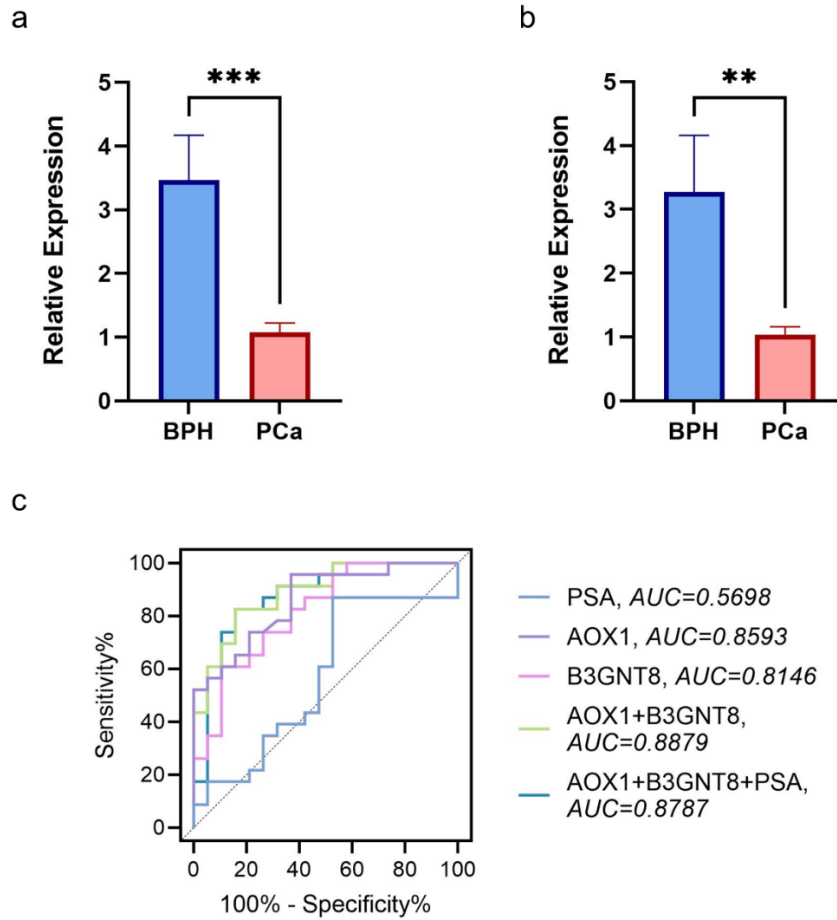


Figure S6. PSA 4-10 ng/mL subgroup analysis: Cancer-specific downregulation of AOX1 (a) and B3GNT8 (b) in PCa patients. (c) ROC curve comparison of liquid biopsy biomarkers' diagnostic performance in PSA “gray zone” patients. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.